Amendments to the Specification:

Please replace the paragraph at page 8, lines 11-25, with the following amended paragraph:

A family of homologous proteins (hereinafter referred to as "Protein Cluster I") was identified by an "all-versus-all" BLAST procedure using all *Caenorhabditis elegans* proteins in the Wormpep20 database release ~~*(http://www.sanger.ac.uk/Projects/C_elegans/wormpep/index.shtml)*~~ *(sanger.ac.uk/Projects/C_elegans/wormpep/index.html)*. The Wormpep database contains the predicted proteins from the C. *elegans* genome sequencing project, carried out jointly by the Sanger Centre in Cambridge, UK and the Genome Sequencing Center in St. Louis, USA. A number of 18,940 proteins were retrieved from Wormpep20. The proteins were used in a Smith-Waterman clustering procedure to group together proteins of similarity (Smith T. F. & Waterman M. S. (1981) *Identification of common molecular subsequences*. J. Mol. Biol. 147(1): 195-197; Pearson W R. (1991) *Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms*. Genomics 11: 635-650; Olsen et al. (1999) *Optimizing Smith-Waterman alignments*. Pac Symp Biocomput.302-313). Completely annotated proteins were filtered out, whereby 10,130 proteins of unknown function could be grouped into 1,800 clusters.

Please replace the paragraph at page 9, lines 1-11, with the following amended paragraph:

The obtained sequence clusters were compared to the *Drosophila melanogaster* proteins contained in the database Flybase (Berkeley Drosophila Genome Project; ~~*http://www.fruitfly.org)*~~ *fruitfly.org*), and annotated clusters were removed. Non-annotated

protein clusters, conserved in both *C. elegans* and *D. melanogaster*, were saved to a worm/fly data set, which was used in a BLAST procedure ~~*(http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/information3.html)*~~ *(ncbi.nlm.nih.gov/Education/BLASTinfo/information3.html)* against the Celera Human Genome Database ~~*(http://www.celera.com)*~~ *(celera.com)*. Overlapping fragments were assembled to, as close as possible, full-length proteins using the PHRAP software, developed at the University of Washington ~~*(http://www.genome.washington.edu/UWGC/analysistools/phrap.htm)*~~ *(genome.washington.edu/UWGC/analysistools/phrap.htm)*. A group of homologous proteins ("Protein Cluster I") with unknown function was chosen for further studies.

Please replace the paragraph at page 9, line 20 – page 10, line 2, with the following amended paragraph:

An alignment of the human polypeptides included in Protein Cluster I (SEQ ID NOS: 2, 4, 6 and 8), using the ClustalX multiple alignment software (downloadable from e.g. ~~ftp://ftp.cbi.ac.uk)~~ cbi.ac.uk) is shown in Table I. For references to the ClustalX software, see Thompson et al. (1997) *The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools*; Nucleic Acids Research, 24:4876-4882. See also Jeanmougin et al. (1998) *Multiple sequence alignment with ClustalX*; Trends Biochem. Sci. 23:403-405. The alignment showed a high degree of conservation in two separate regions, indicating the presence of two novel domains (see positions marked with stars in Table I).

Please replace the paragraph at page 10, lines 4-16, with the following amended paragraph:

A HMM-Pfam search was performed on the human family members. Pfam ~~*http://pfam.wustl.edu*~~ *pfam.wustl.edu* is a large collection of protein families and domains. Pfam contains multiple protein alignments and profile-HMMs (Profile Hidden Markov Models) of these families. Profile-HMMs can be used to do sensitive database searching using statistical descriptions of a sequence family's consensus. Pfam is available on the WWW at ~~*http://pfam.wustl.edu; http://www.sanger.ac.uk/Software/Pfam;* and~~ ~~*http://www.cgr.ki.se/Pfam*~~ *pfam.wustl.edu*; *sanger.ac.uk/Software/Pfam*; and *cgr.ki.se/Pfam*. The latest version (4.3) of Pfam contains 1815 families. These Pfam families match 63% of proteins in SWISS-PROT 37 and TrEMBL 9. For references to Pfam, see Bateman et al. (2000) *The Pfam protein families database*. Nucleic Acids Res. 28:263-266; Sonnhammer et al. (1998) *Pfam: Multiple Sequence Alignments and HMM-Profiles of Protein Domains*. Nucleic Acids Research, 26:322-325; Sonnhammer et al. (1997) *Pfam: a Comprehensive Database of Protein Domain Families Based on Seed Alignments*. Proteins 28:405-420.

Please replace the paragraph at page 10, lines 20-25, with the following amended paragraph:

The human proteins in Cluster I were analyzed using the TM-HMM tool available e.g. at ~~*http://www.cbs.dtu.dk/services/TMHMM-1.0*~~ *cbs.dtu.dk/services/TMHMM-1.0*. TM-HMM is a method to model and predict the location and orientation of alpha helices in membrane-spanning proteins (Sonnhammer et al. (1998) *A hidden Markov model for predicting transmembrane helices in protein sequences*. ISMB 6:175-182). Transmembrane segments were identified in the proteins shown as SEQ ID NOS: 2, 6 and 8 (Fig. 1).

Please replace the paragraph at page 11, lines 13-16, with the following amended paragraph:

The human proteins also show 38% identity to a *Saccharomyces cerevisiae* protein (GenPept Accession No. CAA99495.1). The yeast protein has been annotated in the *Saccharomyces* Genome Database as a putative transporter ~~(http://genome-www4.stanford.edu/cgi-bin/SGD/locus.pl?locus=YOR270c)~~ stanford-edu/cgi-bin/SGD/locus.pl?locus=YOR270c).

Please replace the paragraph at page 11, line 24-page 12, line 2, with the following amended paragraph:

EST databases provided by the EMBL ~~*(http://www.embl.org/Services/index.html)*~~ *(embl.org/Services/index.html)* were used to check whether the human proteins in Cluster I were expressed, in order to identify putative pseudogenes. One putative pseudogenes were identified in Protein Cluster I.

Please replace the paragraph at page 12, lines 3-13, with the following amended paragraph:

The tissue distribution of the human genes was studied using the Incyte ~~LifeSeq®~~ LIFESEQ® database ~~*(http://www.incyte.com)*~~ *(incyte.com)*. The nucleic acid molecule shown as SEQ ID NO: 1 was found to be expressed primarily in the nervous system and the digestive system. The nucleic acid molecule shown as SEQ ID NO: 3 was expressed primarily in male genitalia. The nucleic acid molecule shown as SEQ ID NO: 5 was expressed primarily in the liver and in embryonic structures. The nucleic acid molecule shown as SEQ ID NO: 7 was expressed primarily in the immune system. Therefore, the said nucleic acid molecules shown as SEQ ID NO: 1, 3, 5 and 7 and the polypeptides shown as SEQ ID NO: 2, 4, 6 and 8 are proposed to be useful for differential identification of the

tissue(s) or cell types(s) present in a biological sample and for diagnosis of diseases and disorders, including metabolic disorders and immune diseases.

Please replace the paragraph at page 12, lines 15-21, with the following amended paragraph:

Multiple Tissue Northern blotting (MTN) is performed to make a more thorough analysis of the expression profiles of the proteins in Cluster I. Multiple Tissue Northern (MTN™) Blots ~~*(http://www.clontech.com/mtn)*~~ *(clontech.com/mtn)* are pre-made Northern blots featuring Premium Poly A+ RNA from a variety of different human, mouse, or rat tissues. MTN Blots can be used to analyze size and relative abundance of transcripts in different tissues. MTN Blots can also be used to investigate gene families and alternate splice forms and to assess cross species homology.

Please replace the paragraph at page 13, lines 5-15, with the following amended paragraph:

The expression pattern of the proteins in Cluster I can be analyzed using ~~GeneChip®~~ GENECHIP® expression arrays ~~*(http://www.affymetrix.com/products/app_exp.html)*~~ *(affymetrix.com/products/app_exp.html)*. Briefly, mRNAs are extracted from various tissues. They are reverse transcribed using a T7-tagged oligo-dT primer and double-stranded cDNAs are generated. These cDNAs are then amplified and labeled using In Vitro Transcription (IVT) with T7 RNA polymerase and biotinylated nucleotides. The populations of cRNAs obtained are purified and fragmented by heat to produce a distribution of RNA fragment sizes from approximately 35 to 200 bases. ~~GeneChip®~~ GENECHIP® expression arrays are hybridized with the samples. The arrays are washed and stained. The cartridges are scanned

using a confocal scanner and the images are analyzed with the GeneChip 3.1 software (Affymetrix).

Please replace the paragraph at page 13, line 23-page 14, line 7, with the following amended paragraph:

The two-hybrid method can be used to determine if two known proteins (i.e. proteins for which the corresponding genes have been previously cloned) interact. Another important application of the two-hybrid method is to identify previously unknown proteins that interact with a target protein by screening a two-hybrid library. For reviews, see e.g.: Chien et al. (1991) *The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest*. Proc. Natl. Acad. Sci. U.S.A. 88:9578-9582; Bartel P L, Fields (1995) *Analyzing protein-protein interactions using two-hybrid system*. Methods Enzymol. 254:241-263; or Wallach et al. (1998) *The yeast two-hybrid screening technique and its use in the study of protein-protein interactions in apoptosis*. Curr. Opin. Immunol. 10(2): 131-136. See also ~~*http://www.clontech.com/matchmaker*~~ *clontech.com/matchmaker*.